



变构药物设计与发现研究进展

李睿^{1,2,4}, 何欣恒^{1,4}, 隆兰心¹, 张健^{1,2,3*}

1. 上海交通大学医学院, 药理学与人工智能交叉学系, 上海 200025

2. 上海交通大学医学院附属瑞金医院, 上海交通大学医学院国家转化医学研究中心, 医学基因组学国家重点实验室, 上海 200025

3. 宁夏医科大学药学院, 多肽与蛋白质药物研究中心, 六盘山地区药用资源保护与开发教育部重点实验室, 银川 750004

4. 中国科学院上海药物研究所原创新药研究全国重点实验室, 上海 201203

*通讯作者, E-mail: Jian.zhang@sjtu.edu.cn

收稿日期: 2025-02-28; 接受日期: 2025-04-12; 网络版发表日期: 2025-07-01

国家重点研发计划(编号: 2023YFF1205103)、宁夏回族自治区重点研究与建设计划(编号: 2022BEG01002)、浙江大学上海高级研究院星夜科学基金(编号: SN-ZJU-SIAS-007)和国家自然科学基金(编号: 22237005, 81925034)资助项目

摘要 变构结合位点的识别在结构生物学与计算生物学之间建立了关键联系, 极大地推动了变构药物的发现。然而, 当前变构药物开发的主要策略仍依赖于高通量筛选, 由于对变构效应的理解有限, 这种方法存在较高的失败率。本综述通过收集变构效应案例研究、蛋白质结构数据库和人工智能计算算法的发展, 旨在加深对变构现象的理解, 并指导更有选择性潜力的变构药物开发。这一领域的关键在于将结构生物学与计算生物学相结合, 对于将三维结构数据集转化为可用的药物发现知识至关重要。这些数据集和人工智能算法为建立变构结合位点识别奠定了基础, 从而揭示了变构分子与蛋白的构效关系, 并推动了针对变构蛋白的计算算法的发展, 进而推动了选择性变构药物发现和优化的相关药学研究进步。

关键词 变构效应, 人工智能算法, 药物发现与优化

1 引言

蛋白在维持生命的众多生物功能中扮演着至关重要的角色, 包括能量转移、免疫、生物合成、降解以及细胞通讯^[1-4]。蛋白的三维结构与其配体之间的功能关系仍然是药物设计研究领域中的一个重要研究方向, 被称为构效关系(structure-functional relationship)研究^[5-8]。值得注意的是, 生理状态下的蛋白结构是柔性的, 是一种在体内表现出固有的灵活运动属性的结构。该性质有助于配体与蛋白质内源性活性位点之外的其他氨基酸口袋产生相互作用, 影响蛋白的结构变化, 从

而产生变构调节作用(激动或者抑制蛋白功能), 这种现象被称为变构(Allostery)^[9]。从配体结合蛋白口袋的结构分类来说, 内源性配体常常结合在蛋白靶点的内源性口袋(正构)位点, 而靶点的内源性正构(Orthosteric)口袋具有高度的进化保守性。因此针对这些保守(正构)位点的药物化合物常常面临靶点选择性差, 或多靶点导致副作用多等问题。变构调节提供了一种高专一性潜力的调节蛋白功能的策略。与直接抑制或激活蛋白的活性位点不同, 变构调节涉及调节分子在活性位点之外的位点(变构位点)上的相互作用, 这些位点通常在空间上与酶活性或内源性配体结合的位置不

引用格式: Li R, He X, Long L, Zhang J. Research progress on design and discovery of allosteric drugs. *Sci Sin Chim*, 2025, 55: 2286–2295, doi: 10.1360/SSC-2025-0046

同。因此变构调节剂不会直接与活性位点的内源性底物竞争。结合这些变构位点通常也会导致蛋白结构的构象变化, 这些变化可以增强或抑制蛋白的活性, 从而调节其功能。这种变构调节策略为设计相关疾病蛋白的调节剂分子, 从而进行潜在专一性治疗干预提供了一种有价值并且副作用较少的途径。

变构位点固有的专一性潜力在选择性调节蛋白功能的药物设计中起着关键作用, 有助于高精度区分不同亚型的同家族蛋白靶点, 从而实现更精细的蛋白家族亚型的靶向蛋白精准治疗和干预, 从而最小化对蛋白基本生理功能的干扰, 减少发生不良反应的可能。变构位点的发现和利用为推进药理学更进一步地发展带来了新的希望, 特别是考虑到这些变构调节现象在一些功能关键的蛋白(如G蛋白偶联受体(GPCRs)、酶和离子通道)中广泛存在。这进一步强调了在变构位点识别和药物设计中纳入蛋白动态变构结构见解的必要性, 如图1所示。通过动力学等计算方法研究变构口袋和变构效应对于专一性变构药物的发现和设计的持续创新至关重要, 促进了高效且低毒的新型治疗药物的开发。例如, 对GPCRs、Kirsten大鼠肉瘤病毒癌基因同源物(KRAS)和表皮生长因子受体(EGFR)^[10-13]等蛋白中变构结合位点的研究展示了这种潜力, 并凸显了

如何利用对蛋白结构的理解来开发变构调节剂, 并绕过与直接竞争和脱靶效应相关的挑战^[14]。

变构理论的框架自1904年被克里斯蒂安·玻尔(Christian Bohr)^[15]发现以来得到了显著发展, 其通过研究血红蛋白中氧气分子或二氧化碳分子的自调节结合现象, 揭示了同类型分子结合亲和力的增强机制。1961年, “变构”(allosteric)这一术语被正式提出, 用以描述蛋白质通过活性位点以外的调控位点发生构象调整从而实现功能调控的过程。该理论最初通过两个经典模型得以阐释, 分别为20世纪60年代提出的Koshland、Nemethy和Filmer (KNF)模型, 以及Monod、Wyman和Changeux (MWC)协同模型^[16,17]。后续研究持续推进了理论发展。1984年, Cooper和Dryden^[18]提出了动态变构模型, 强调构象动力学在变构调控中的流动性特征。1999年, “群体漂移与构象选择”理论框架的建立进一步深化了对变构效应的理解^[19,20]。2004年, Nussinov团队^[21]提出了变构作为蛋白质的固有属性这一突破性观点, 揭示了先进变构疗法在重大疾病治疗中的重要作用。同年, 首个变构药物获美国食品药品监督管理局批准上市。至2015年, 该理论体系进一步扩展, 明确了变构效应在蛋白质系统中具有普适性特征^[22]。

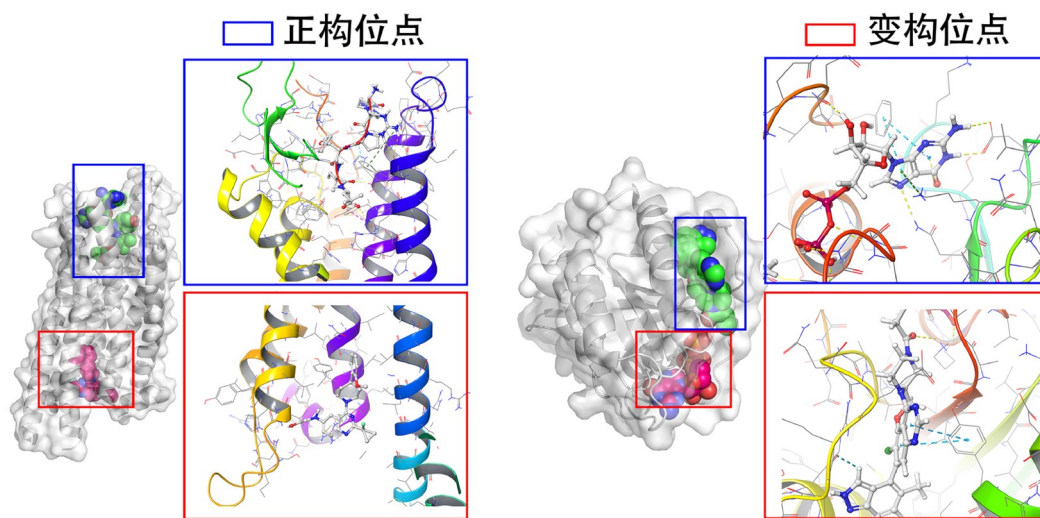


图 1 (网络版彩图)蛋白质-配体复合物的实验证据揭示了变构位点和正构位点, 包括NTSR1在NTSR1-GRK2-Gα复合物中的冷冻电镜结构(PDB: 8JPF, 分辨率: 3.02 Å, 左图)和KRASG12C配体复合物的X射线结构(PDB: 7O83, 分辨率: 2.38 Å, 右图)

Figure 1 (Color online) Experimental evidence from protein-ligand complexes has revealed allosteric and orthosteric sites, including the cryo-electron microscopy structure of the NTSR1-GRK2-Gα complex (PDB: 8JPF, resolution: 3.02 Å, left) and the X-ray structure of the KRASG12C ligand complex (PDB: 7O83, resolution: 2.38 Å, right).

近几年, 结构生物学领域的最新突破为变构理论带来了深刻见解. 其中, 低温电子显微镜(Cryo-EM)的贡献尤为显著^[23,24]. 同时, 核磁共振(NMR)作为一种本质上为动态的方法, 也对我们理解变构蛋白机制作出了重要贡献^[25~27], 并且可以用于表征变构结合位点处于“开放”或“关闭”状态的概率, 正如最近对PDZ2结构域的研究结果所示^[27]. 此外, 各种实验技术, 如点突变与生物活性测定相结合, 已被应用于识别变构结合位点^[28,29]. 将人工智能(AI)和机器学习(ML)应用于变构位点识别中的创新方法, 为调节通路激活机制提供了新的见解^[30,31]. 这些方法不仅阐明了突变对各种效应物的抑制作用, 还揭示了结合特异性的变化, 为治疗干预提供了新的途径.

在本综述中, 我们收集并讨论了与变构药物的开发和发现相关的最新结构进展、结构数据库的创建和计算方法(图2), 对其使用进行了举例说明, 讨论了支持这些内容的ML技术, 并强调了当前用于理解变构在药物设计中作用的工作流程.

2 变构调控相关的实验证据与案例

揭示变构调控事件相关的结构解析实验证据, 有助于进一步证明变构调控的真实性和普遍性^[25,32]. 这为使用变构效应进行药物设计所涵盖的蛋白范围提供更好的证据. 有研究使用冷冻电镜技术系统地描述了变构调节剂对 μ -阿片受体(MOR)的影响, 阐明了配体如何通过GPCRs核心区域的变构口袋结构诱导GPCR蛋白的激活^[23,29]. 此外, 对人类酪氨酸磷酸酶蛋白PDZ结构域的变构结构研究绘制了变构相互作用的网络, 揭示了激活机制^[27,33,34]. 另外, 对KRAS变构氨基酸通

讯的研究揭示了氨基酸突变如何改变三维空间结构中的自由能景观, 进而影响蛋白的结合口袋的氨基酸特异性. 总之, 使用最新的结构解析技术方法已经展现并强调了蛋白变构在药物发现和开发中的广泛适用性和有效性, 为多种不同靶点蛋白的治疗和相关疾病的干预提供了深入见解.

2.1 通过冷冻电镜策略研究GPCR蛋白的变构调节

GPCRs在变构调节剂触发下的变构信号转导机制尚未被完全阐明^[24]. μ -阿片受体(MOR)作为A类GPCR家族的一员, 因其在镇痛中的作用及其对吗啡分子的特异性结合亲和力, 在公共卫生领域具有重要意义.

近年来, 冷冻电镜技术的发展在阐明MOR (以及其他GPCR变构研究, 如GPR34等^[24,35])的结构变化方面发挥了关键作用. 特别是, 利用变构调节分子BMS-986122的实验生成了两种不同的密度图, 揭示了在调节剂存在和不存在情况下MOR (DAMGO)-Gi-scFv16复合物的三维结构^[23,29]. 以上研究强调了保守残基R167^{3.50}和Y254^{5.58}的关键作用. 其氢键相互作用在稳定外移的跨膜螺旋6 (TM6)中起到了重要作用. 此外, BMS-986122与跨膜螺旋3 (TM3)的结合改变了其与跨膜螺旋5 (TM5)的状态, 从而通过变构影响促进了保守残基R167^{3.50}和Y254^{5.58}之间关键相互作用的稳定. 这种相互作用的变构影响使得跨膜区域成为增强GPCR活性的潜在靶点, 为新药发现和开发提供了新的见解. 未来可以将更多的冷冻电镜研究中的结构见解整合到GPCR靶向治疗变构药物的设计中, 为开发新型镇痛药及其他针对MOR和相关受体的药物提供具有前景的研究思路. 对变构位点残基的研究加速了针对GPCR介导疾病的有效治疗药物的合理开发.

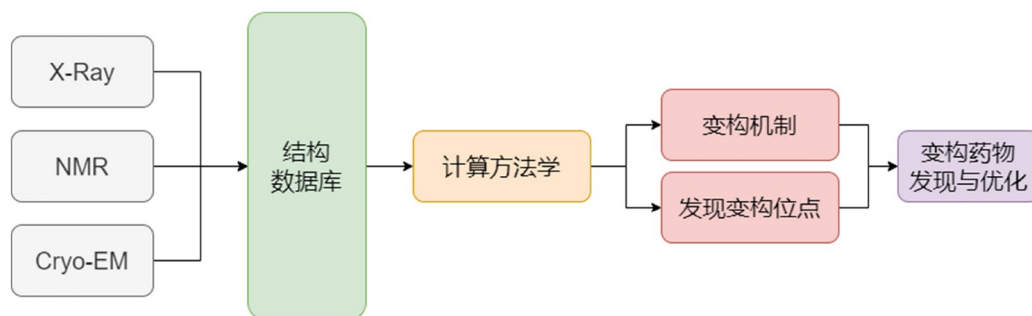


图2 (网络版彩图)本综述的研究工作流程

Figure 2 (Color online) The research flowchart of this review.

2.2 通过NMR研究PDZ结构域蛋白的变构调节

近年来, 溶液NMR技术取得的相关进展也促进了多状态的蛋白结构解析, 特别是对人类酪氨酸磷酸酶蛋白的PDZ结构域领域的研究^[30]. 这项研究为该蛋白的变构位点及相关结构差异提供了宝贵的见解, 解释了变构激活机制, 对全面理解结构变化如何影响蛋白质功能(如识别、生物信号转导和降解)至关重要. 研究主要解释了蛋白结构内动态变构的特征模式. 在原子级别分辨率下, 基于化学位移和弛豫的NMR研究提供了蛋白在水溶液缓冲液中折叠时的氨基酸间变构交互的详细分析. 以上研究允许量化和可视化地研究配体结合所引起的氨基酸化学位移, 从而提供变构位点的结构信息^[27]. 这些位移既可以是局部的(围绕结合位点, 伴随直接构象变化), 也可以是全局的(远离配体结合位点), 从而全面展示蛋白的敏感变构影响.

此外, NMR结果在定义蛋白结构中残基的相互作用网络从而揭示变构氨基酸路径方面发挥了重要作用^[25,26,36]. 这些氨基酸网络不仅以其动态特性为特征, 还以其能量特性为特征. NMR结果提供了从超过10 ns到微秒甚至毫秒时间尺度上状态间结构相关网络的直接测量证据. 蛋白结构不同结构域的相似比率表明, 不同结构之间存在相关的动态网络联系. 在PDZ2结构域中, 多状态的蛋白结构解析能够有效捕捉蛋白变构的动态结构特征, 补充了在多个维度上的数据, 进一步支持了变构效应在天然蛋白结构中普遍存在的研究理论.

2.3 通过深度突变扫描研究KRAS蛋白的变构调节

KRAS蛋白在体内作为分子开关具有两种状态: 非活性GDP结合状态和活性GTP结合状态. 与激酶以及MOR蛋白类似, KRAS也表现出变构特性, 其特征为结构中氨基酸相互作用信息的长程传递, 随蛋白激动/抑制状态变化而变化. 最近, 一种名为Sotorasib的抑制剂已获准进行临床试验^[37]. 该化合物作为KRAS (G12C)突变的共价结合剂, 能够专一性地在核苷酸和效应子结合位点外通过变构方式抑制KRAS (G12C), 从而将其锁定在非活性GDP结合状态^[12]. 该研究包含22000多个自由能测量值, 揭示了致癌KRAS的突变效应与野生型KRAS中的突变效应高度相关(例如, 在人类GAP共表达式存在的情况下, 野生型KRAS和KRAS

(G12C)之间 $r = 0.93$, $n = 776$). 然而, 对KRAS变构结合位点理解的缺乏仍然是针对KRAS蛋白的其他药物发现和开发工作的障碍.

通过将表面等离子体共振(SPR)实验与KRAS氨基酸点突变实验相结合, 有研究生成了广泛的构效关系图谱, 展示了KRAS蛋白、gC1qR/C1QBP/P32蛋白和糖皮质激素受体等的抑制性变构位点^[12]. 具体而言, 该方法详细比较了氨基酸点突变如何影响KRAS相互作用的动力学能量图景. 这项研究成功量化了野生型KRAS的氨基酸点突变的影响, 从而丰富了对KRAS变构事件的实验机制理解. 该方法为探测蛋白结构中未识别的变构位点提供了可以参考的实验框架, 并为可能在将来通过触发蛋白变构理解和调节其他蛋白活性的方法开发提供研究基础.

3 变构相关计算方法与药物设计案例

不同的蛋白相关实验证据表明, 蛋白质的三维结构可以为蛋白变构和药物设计提供关键机制及氨基酸相互作用位点等重要信息. 这些信息的广泛应用促使研究人员进行了大量的蛋白三维空间结构数据的收集, 期望进一步发展用于蛋白变构位点预测等的计算方法及相关研究, 从而进一步促进具有专一性潜力的变构药物的发现方法的开发. 因此, 本文总结了一些重要的高质量蛋白三维空间结构数据集(表1^[38-48]), 这对计算蛋白变构位点方法的开发和变构药物设计的探索至关重要.

3.1 变构蛋白三维结构数据库

结构生物学^[38]在变构研究中提供了关键见解, 包括对变构位点的识别和变构残基相互作用网络^[39]的绘制. 这些数据对于开发用于理解变构及其在药物发现中的意义的计算工具至关重要^[40,41]. 由此建立的实验数据集为针对变构蛋白的计算工具奠定了坚实基础, 提供了对变构药物发现和开发至关重要的详细结构见解.

尽管取得了以上进展, 但结构生物学中的实验方法并非没有局限性, 通常需要大量的资金和时间投入, 并且仍然存在失败的风险. 计算方法有可能加速临床前药物研究, 并通过以较低成本提供预测分析来补充结构生物学, 从而预测具有生物活性功能的关键残

表 1 目前的结构数据库和计算工具研究进展

Table 1 The current research progress in structural databases and computational tools

算法	功能	网址或工具	适用范围	参考文献
Allosteric Database 2023		http://mdl.shsmu.edu.cn/ASD/	变构蛋白	[38]
AFDB	结构数据库	https://alphafold.ebi.ac.uk/	(预测)蛋白结构	[39]
AlloMAPS2		http://allomaps.bii.a-star.edu.sg	(实验)蛋白结构	[40]
PrankWeb		3D CNN	CHEN11数据集, JOINED数据集, COACH420数据库和HOLO4K数据集	[41]
PocketMiner		GNN	39个口袋示例的测试数据集	[42]
AlloReverse		AdaBoost	58个蛋白质	[43]
CorrSite	变构位点发现	GVP	35个变构蛋白	[44]
TopoAlloSite		SVM	多个关键变构蛋白, 包括磷酸泛酰半胱氨酸合成酶、亚精胺合成酶和Sirtuin 6系统	[45]
PASSerRank		LambdaMART	ASD	[46]
PASSer2.0		Ensemble	基于FPocket识别的口袋特征	[47]
PASSer		XGBoost + GCNN	ASD、ASBench和CASBench	[48]

基^[42-44].

为进一步增强对变构蛋白结构的理解, 研究人员从论文、专利和其他公开来源收集了与变构研究相关的信息. 变构数据库(ASD, 2023版)收集了66589个潜在变构位点、748个变构蛋白-蛋白相互作用(PPI)调节剂、87个明确变构中6565个先导化合物和456个双重调节剂的数据. 这些双重调节剂同时调节变构和正构位点^[45,46]. 这一收集标志着基础变构数据的显著增长. AlloMAPS2数据集的更新包含变构信号图(ASMs)和变构探测图(APMs), 量化了突变的变构效应^[47]. 该数据集包含46个具有变构位点的蛋白质、1908个蛋白质链和33个蛋白质, 每个蛋白质具有超过50个已知的病理SNP. 这使得该数据集能够在发现潜在调控位点和设计变构位点及效应子时, 准确评估变构信号效应, 从而发挥一定作用. 此外, AlphaFold数据库(AFDB)包含超过2亿个结构位点^[48].

3.2 机器学习方法在变构结合位点识别中的应用

变构涉及一种调节机制, 其中一个分子在蛋白质的一个位点结合会影响远处残基位点的活性, 这种影响通过长程分子内通讯实现^[11,14,49]. 识别这些远程变构位点可以揭示用于虚拟筛选变构拮抗剂和激动剂的潜在口袋, 从而为变构药物发现作出重要贡献^[11,50]. 为扩展对潜在变构结合位点的了解, 避免高成本的实验检测, 各种计算方法已被开发并应用于三维蛋白质结

构数据库或单个结构. 通过结合基于物理的实验结果和机器学习方法, 这些计算策略在解决变构位点识别问题方面发挥了重要作用. 本节描述了利用变构结构信息预测变构结合位点的代表性机器学习工具, 并重点介绍了它们的方法、应用及其在药物发现领域的影响.

2023年, Xie等^[51]开发了TopoAlloSite方法, 使用支持向量机(support vector mechine, SVM)模型和子图核匹配, 采用以CATH结构域为节点的蛋白质结构网络表示变构位点, 同时进一步研究说明, 变构位点(包括隐秘位点)的位置取决于蛋白质拓扑结构. 该方法的召回率为0.71, 准确率为0.76, 能够有效优先考虑各种蛋白质中的真实变构位点. 该方法准确预测了多个关键变构蛋白中先前未被识别的变构位点, 包括磷酸泛酰半胱氨酸合成酶、亚精胺合成酶和Sirtuin 6系统^[52]. 上述结果凸显了TopoAlloSite方法在识别功能相关变构位点方面的显著效用, 这些位点可能被传统结构分析所掩盖. 同年, PASSerRank利用LambdaMART模型, 这是一种结合梯度提升决策树和LambdaRank函数的集成学习方法, 用于根据变构位点的可能性进行排名预测^[53]. 该方法在ASD中的F1得分为0.662, 在CAS-Bench数据库中的F1得分为0.608, 同时在ASD和CAS-Bench中的马修斯相关系数(MCC)分别为0.645和0.589.

此外, 神经网络及其增强方法也在该领域得到了

应用。PASSer结合了XGBoost模型和图卷积神经网络(GCNN),使用口袋的原子网络表示变构位点,并基于FPocket物理化学特征进行训练,以增强预测能力^[54]。该方法在多个基准测试中表现出色,在ASD、AS-Bench和CASBench中的F1得分均为0.782。此外,其Top 3正确率为0.849,进一步验证了其在准确识别变构位点方面的有效性。

与此同时,其他机器学习方法如随机森林(RF)等也被应用于变构位点识别,使用专门工具进行动态和拓扑特征分析。ESSA方法专注于分析动态和拓扑特征,而不采用机器学习,为变构位点识别提供了统计标准^[55]。在与PARS和AllositePro的比较分析中,ESSA方法展示了更高的预测能力。具体而言,其在数据集II中成功预测了14个变构位点中的10个,成功率高于PARS(预测2/14)和AllositePro(预测8/14)。PrankWeb采用了一种策略,识别3D网格上的可配体点并将其聚类为表面斑块作为预测结合位点^[56]。该模型使用CHEN11数据集进行训练,并使用JOINED数据集进行验证。在测试中,该模型使用了COACH420数据库和HOLO4K数据集,Top-*N*成功率分别为0.720和0.686。BiteNet将3D蛋白质构象视为图像,将变构位点检测视为图像中的对象检测任务^[57]。该方法的预测能力在COACH420和HOLO4K数据集中得到了评估,其Top-*N*成功率分别超过0.6和0.7。CorrSite作为CavityPlus平台的一部分,通过结构运动特征和生物物理特性揭示了正构位点和变构位点之间的耦合^[58]。CorrSite2.0展示了高预测准确性,成功识别了数据集I中35个变构蛋白中36个变构位点中的35个。这些蛋白质的序列长度从169到917个氨基酸不等。此外,在使用CorrSite2.0数据集II的独立测试阶段,该模型准确预测了20个变构位点中的18个。此外,PASSer2.0在PASSer的基础上,使用AutoGluon训练了一组模型,基于FPocket识别的相同口袋特征,专注于方法改进而不进行直接比较分析^[59]。该方法的F1得分为0.701,Top 3位置准确率为0.827。

在增强采样和隐秘口袋方面,PocketMiner结合了两种方法:分子动力学(MD)采样与机器学习算法,以预测隐秘口袋,揭示潜在的变构位点^[60]。该算法的训练数据集包括分子动力学模拟,这些模拟识别了SARS-CoV-2蛋白质组中大多数蛋白质的隐秘口袋、先前模拟的埃博拉病毒VP35蛋白数据以及一组16个

已知配体结合隐秘口袋的蛋白质。这些蛋白质专门为本研究进行了模拟。为评估PocketMiner的预测能力,从PDB中策划了一个包含39个口袋示例的测试数据集,使用0.7的阈值,PocketMiner的总体准确率为0.90,成功识别了4个阴性示例中的3个和24个阳性示例中的21个。KeyAlloSite使用进化耦合模型预测变构位点并识别关键变构残基。该方法通过成对比较变构口袋内每个残基与功能位点残基之间的进化耦合分数来推断关键变构残基^[61],已在典型变构蛋白(如BCR-ABL1、Tar和PDZ3)的先前实验研究中得到验证。AlloReverse使用无偏MD模拟进行蛋白质holo和apo状态之间的比较自由能分析,通过能量耦合分析识别潜在的变构位点^[62]。该模型使用了包含58个蛋白质的测试集进行基准测试,包括926个口袋样区域,其中83个被注释为“变构位点”。该模型的召回率为71.0%,接收者操作特征(ROC)曲线下面积(AUC)为0.758,表明该模型在复杂数据集中具备有效识别变构位点的能力。Pillai等^[63]基于MWC模型,开发了一种从头设计算法,用于生成多种可切换的变构系统(图3)。

预测变构结合位点的计算框架从传统方法(如SVM与FPocket的结合)显著发展到了利用AI的多种独特策略。本综述强调了向AI驱动方法的转变。这些方法不仅增强了这些框架的效能,还引入了一系列创新方法,在预测准确性和计算效率方面提供了显著优势。此外,AI算法的常规评估涉及F1分数、召回率、ROC曲线下面积和Top-*N*准确率等指标。然而,由于每个算法使用的数据集不同,这些性能指标在不同研究之间

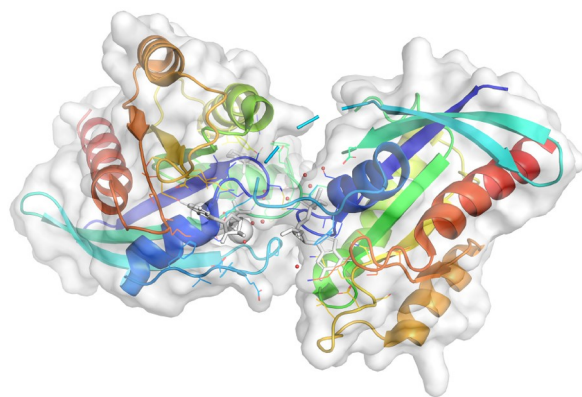


图3 (网络版彩图) KRASG12C蛋白的溶液核磁结构
Figure 3 (Color online) The solution NMR structure of KRASG12C protein.

不能进行直接比较。这些进展凸显了AI在增强变构蛋白位点计算分析中的关键作用, 并为该领域的未来探索描绘了有前景的轨迹。

4 总结与展望

变构效应是非常重要的选择性药物设计策略。借鉴变构策略进行药物设计的一个经典案例是使用DNA分子的可编程策略构筑双价配体分子, 用于开发“正构-变构”共同作用的双价配体分子库, 实现了变构-正构双靶点协同策略, 从而高效筛选获得对受体识别亲和力最高、激动效能最强的双价配体分子。该策略为双价配体分子的设计开发提供了全新的思路, 在药物设计及生物医学研究领域具有广泛的应用前景。与此同时, 模型的可解释性也需要进一步设计与研究。例如, SHAP值和特征重要性排名可以用于分析机器学习模型如何利用蛋白质的结构和动态特征来预测变构位点, 揭示模型在识别变构位点时依赖的关键特征, 不仅可以提高AI模型在变构位点识别中的透明度和可信度, 还可以为药物发现提供有价值的见解, 帮助研究人员更好地理解变构调节的分子机制。总而言之, 结合位点的识别代表了结构与计算之间的重要联系, 显著增强了变构药物的发现。这些药物具有特异性、非竞争性和可调性。尽管取得了以上进展, 但变构药物的开发策略^[64,65](图4和图5)主要依赖于高通量筛选, 其高失败率主要归因于对变构效应的不完全理解。本综述旨在整合变构效应案例研究、蛋白质结构数据库和算法的见解, 以增强对变构效应的理解, 从而为更有效的变构药物开发提供信息, 并加深对蛋白质变构的理解。结构生物学与计算方法的结合在变构药物的发现研究中至关重要, 必须认识到包含广泛三维结构数据(包括柔性结构数据)的高质量数据集的必要性。这

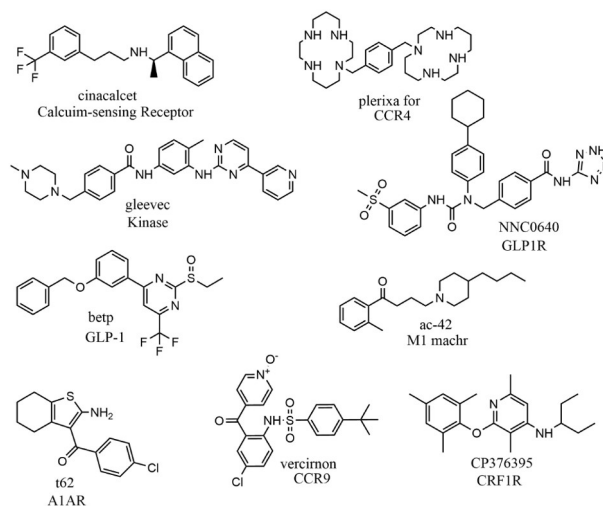


图4 变构分子及其靶标(I)

Figure 4 Allosteric molecules and their targets (I).

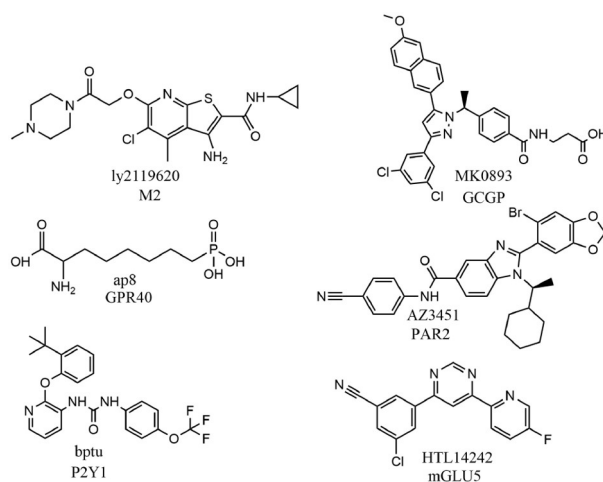


图5 变构分子及其靶标(II)

Figure 5 Allosteric molecules and their targets (II).

些数据集对于建立结构-活性关系(SARs)至关重要, 并且是未来开发变构蛋白计算算法的核心。

参考文献

- 1 Tesei G, Trolle AI, Jonsson N, Betz J, Knudsen FE, Pesce F, Johansson KE, Lindorff-Larsen K. *Nature*, 2024, 626: 897–904
- 2 Kawai A, Tokunoh N, Kawahara E, Tamiya S, Okamura S, Ono C, Anindita J, Tanaka H, Akita H, Yamasaki S, Kunisawa J, Okamoto T, Matsuura Y, Hirai T, Yoshioka Y. *J Clin Invest*, 2023, 133: e166827
- 3 Zhao L, Zhao J, Zhong K, Tong A, Jia D. *Sig Transduct Target Ther*, 2022, 7: 113
- 4 Tokizane K, Brace CS, Imai S. *Cell Metab*, 2024, 36: 377–392.e11
- 5 Li Z, Wu Q, Huang G, Jin X, Li J, Pan X, Yan N. *Proc Natl Acad Sci USA*, 2024, 121: e2322899121

- 6 Guo Q, He B, Zhong Y, Jiao H, Ren Y, Wang Q, Ge Q, Gao Y, Liu X, Du Y, Hu H, Tao Y. *Nat Chem Biol*, 2024, 20: 74–82
- 7 Susa KJ, Kruse AC, Blacklow SC. *Trends Cell Biol*, 2024, 34: 509–522
- 8 Li J, Zhang J, Chen J, Luo X, Zhu W, Shen J, Liu H, Shen X, Jiang H. *J Comb Chem*, 2006, 8: 326–337
- 9 Campitelli P, Modi T, Kumar S, Ozkan SB. *Annu Rev Biophys*, 2020, 49: 267–288
- 10 Hicks KG, Cluntun AA, Schubert HL, Hackett SR, Berg JA, Leonard PG, Ajalla Aleixo MA, Zhou Y, Bott AJ, Salvatore SR, Chang F, Blevins A, Barta P, Tilley S, Leifer A, Guzman A, Arok A, Fogarty S, Winter JM, Ahn HC, Allen KN, Block S, Cardoso IA, Ding J, Dreveny I, Gasper WC, Ho Q, Matsuura A, Palladino MJ, Prajapati S, Sun P, Tittmann K, Tolan DR, Unterlass J, VanDemark AP, Vander Heiden MG, Webb BA, Yun CH, Zhao P, Wang B, Schopfer FJ, Hill CP, Nonato MC, Muller FL, Cox JE, Rutter J. *Science*, 2023, 379: 996–1003
- 11 Krishna Kumar K, O'Brien ES, Habrian CH, Latorraca NR, Wang H, Tuneew I, Montabana E, Marqusee S, Hilger D, Isacoff EY, Mathiesen JM, Kobilka BK. *Cell*, 2023, 186: 1465–1477.e18
- 12 Weng C, Faure AJ, Escobedo A, Lehner B. *Nature*, 2024, 626: 643–652
- 13 Wang Y, Ji D, Lei C, Chen Y, Qiu Y, Li X, Li M, Ni D, Pu J, Zhang J, Fu Q, Liu Y, Lu S. *Comput Struct Biotechnol J*, 2021, 19: 1184–1199
- 14 Abhishek S, Deeksha W, Nethravathi KR, Davari MD, Rajakumara E. *Comput Struct Biotechnol J*, 2023, 21: 5003–5015
- 15 Bohr C, Hasselbalch K, Krogh A. *Skand Arch Physiol*, 1904, 16: 402–412
- 16 Monod J, Wyman J, Changeux JP. *J Mol Biol*, 1965, 12: 88–118
- 17 Koshland Jr DE, Némethy G, Filmer D. *Biochemistry*, 1966, 5: 365–385
- 18 Cooper A, Dryden DTF. *Eur Biophys J*, 1984, 11: 103–109
- 19 Ma B, Kumar S, Tsai CJ, Nussinov R. *Protein Eng Des Sel*, 1999, 12: 713–720
- 20 Kumar S, Ma B, Tsai CJ, Wolfson H, Nussinov R. *Cell Biochem Biophys*, 1999, 31: 141–164
- 21 Gunasekaran K, Ma B, Nussinov R. *Proteins*, 2004, 57: 433–443
- 22 Macpherson JA, Anastasiou D. *Curr Opin Biotechnol*, 2017, 48: 102–110
- 23 Kaneko S, Imai S, Uchikubo-Kamo T, Hisano T, Asao N, Shirouzu M, Shimada I. *Nat Commun*, 2024, 15: 3544
- 24 Duan J, He XH, Li SJ, Xu HE. *Nat Rev Endocrinol*, 2024, 20: 349–365
- 25 Wayment-Steele HK, Ojoawo A, Otten R, Apitz JM, Pitsawong W, Hömberger M, Ovchinnikov S, Colwell L, Kern D. *Nature*, 2024, 625: 832–839
- 26 Saravanan V, Ahammed I, Bhattacharya A, Bhattacharya S. *J Biomol Structure Dyn*, 2024, 42: 1812–1825
- 27 Ashkinadze D, Kadavath H, Pokharna A, Chi CN, Friedmann M, Strotz D, Kumari P, Mingos M, Cadalbert R, Königl S, Güntert P, Vögeli B, Riek R. *Nat Commun*, 2022, 13: 6232
- 28 Maschietto F, Morzan UN, Tofoleanu F, Gheeraert A, Chaudhuri A, Kyro GW, Nekrasov P, Brooks B, Loria JP, Rivalta I, Batista VS. *Nat Commun*, 2023, 14: 2239
- 29 Kaneko S, Imai S, Asao N, Kofuku Y, Ueda T, Shimada I. *Proc Natl Acad Sci USA*, 2022, 119: e2121918119
- 30 Wu N, Barahona M, Yaliraki SN. *Curr Opin Struct Biol*, 2024, 84: 102737
- 31 Nerín-Fonz F, Cournia Z. *Curr Opin Struct Biol*, 2024, 85: 102774
- 32 Hekstra DR. *Annu Rev Biophys*, 2023, 52: 255–274
- 33 Faure AJ, Domingo J, Schmiedel JM, Hidalgo-Carcedo C, Diss G, Lehner B. *Nature*, 2022, 604: 175–183
- 34 Rehman AU, Lu S, Khan AA, Khurshid B, Rasheed S, Wadood A, Zhang J. *Expert Opin Drug Discov*, 2022, 17: 283–295
- 35 Xia A, Yong X, Zhang C, Lin G, Jia G, Zhao C, Wang X, Hao Y, Wang Y, Zhou P, Yang X, Deng Y, Wu C, Chen Y, Zhu J, Tang X, Liu J, Zhang S, Zhang J, Xu Z, Hu Q, Zhao J, Yue Y, Yan W, Su Z, Wei Y, Zhou R, Dong H, Shao Z, Yang S. *Proc Natl Acad Sci USA*, 2023, 120: e2308435120
- 36 Berezhovsky IN, Nussinov R. *J Mol Biol*, 2022, 434: 167751
- 37 Fakih MG, Salvatore L, Esaki T, Modest DP, Lopez-Bravo DP, Taieb J, Karamouzis MV, Ruiz-Garcia E, Kim TW, Kuboki Y, Meriggi F, Cunningham D, Yeh KH, Chan E, Chao J, Saportas Y, Tran Q, Cremolini C, Pietrantonio F. *N Engl J Med*, 2023, 389: 2125–2139
- 38 Garcia-Nafria J, Tate CG. *Biochem Soc Trans*, 2021, 49: 2345–2355
- 39 Kozlovskii I, Popov P. *NAR Genomics BiolInf*, 2021, 3: lqab111
- 40 Dietrich JD, Longenecker KL, Wilson NS, Goess C, Panchal SC, Swann SL, Petros AM, Hobson AD, Ihle D, Song D, Richardson P, Comess KM, Cox PB, Dombrowski A, Sarris K, Donnelly-Roberts DL, Duignan DB, Gontsyan A, Jung P, Krueger AC, Mathieu S, McClure A, Stoll

- VS, Wetter J, Mankovich JA, Hajduk PJ, Vasudevan A, Stoffel RH, Sun C. *J Med Chem*, 2021, 64: 417–429
- 41 Tastan Bishop Ö, Musyoka TM, Barozi V. *J Mol Biol*, 2022, 434: 167610
- 42 Koehler Leman J, Szczerbiak P, Renfrew PD, Gligorijevic V, Berenberg D, Vatanen T, Taylor BC, Chandler C, Janssen S, Pataki A, Carrierio N, Fisk I, Xavier RJ, Knight R, Bonneau R, Kosciolk T. *Nat Commun*, 2023, 14: 2351
- 43 Cheng J, Novati G, Pan J, Bycroft C, Žemgulytė A, Applebaum T, Pritzel A, Wong LH, Zielinski M, Sargeant T, Schneider RG, Senior AW, Jumper J, Hassabis D, Kohli P, Avsec Ž. *Science*, 2023, 381: eadg7492
- 44 Barrio-Hernandez I, Yeo J, Jänes J, Mirdita M, Gilchrist CLM, Wein T, Varadi M, Velankar S, Beltrao P, Steinegger M. *Nature*, 2023, 622: 637–645
- 45 He J, Liu X, Zhu C, Zha J, Li Q, Zhao M, Wei J, Li M, Wu C, Wang J, Jiao Y, Ning S, Zhou J, Hong Y, Liu Y, He H, Zhang M, Chen F, Li Y, He X, Wu J, Lu S, Song K, Lu X, Zhang J. *Nucleic Acids Res*, 2024, 52: D376–D383
- 46 Liu X, Lu S, Song K, Shen Q, Ni D, Li Q, He X, Zhang H, Wang Q, Chen Y, Li X, Wu J, Sheng C, Chen G, Liu Y, Lu X, Zhang J. *Nucleic Acids Res*, 2019, 48: D394–D401
- 47 Tan ZW, Tee WV, Guarnera E, Berezovsky IN. *Nucleic Acids Res*, 2023, 51: D345–D351
- 48 Varadi M, Anyango S, Deshpande M, Nair S, Natassia C, Yordanova G, Yuan D, Stroe O, Wood G, Laydon A, Židek A, Green T, Tunyasuvunakool K, Petersen S, Jumper J, Clancy E, Green R, Vora A, Lutfi M, Figurnov M, Cowie A, Hobbs N, Kohli P, Kleywegt G, Birney E, Hassabis D, Velankar S. *Nucleic Acids Res*, 2022, 50: D439–D444
- 49 Bera A, Mukherjee S, Patra N. *Phys Chem Chem Phys*, 2024, 26: 17011–17027
- 50 Shankar S, Ramachandran S, Tulsian N, Radhakrishnan S, Jobichen C, Sivaraman J. *FEBS J*, 2023, 290: 2437–2448
- 51 Xie J, Pan G, Li Y, Lai L. *J Chem Phys*, 2023, 158: 105102
- 52 Lu S, Chen Y, Wei J, Zhao M, Ni D, He X, Zhang J. *Acta Pharm Sin B*, 2021, 11: 1355–1361
- 53 Tian H, Xiao S, Jiang X, Tao P. *J Comput Chem*, 2023, 44: 2223–2229
- 54 Tian H, Xiao S, Jiang X, Tao P. *Nucleic Acids Res*, 2023, 51: W427–W431
- 55 Kaynak BT, Bahar I, Doruker P. *Comput Struct Biotechnol J*, 2020, 18: 1577–1586
- 56 Jakubec D, Skoda P, Krivak R, Novotny M, Hoksza D. *Nucleic Acids Res*, 2022, 50: W593–W597
- 57 Kozlovskii I, Popov P. *Commun Biol*, 2020, 3: 618
- 58 Xie J, Wang S, Xu Y, Deng M, Lai L. *J Chem Inf Model*, 2022, 62: 187–195
- 59 Xiao S, Tian H, Tao P. *Front Mol Biosci*, 2022, 9: 879251
- 60 Meller A, Ward M, Borowsky J, Lotthammer JM, Kshirsagar M, Oviedo F, Ferres JL, Bowman GR, *bioRxiv*, 2022, <https://doi.org/10.1101/2022.06.28.497399>
- 61 Xie J, Zhang W, Zhu X, Deng M, Lai L. *eLife*, 2023, 12: e81850
- 62 Zha J, Li Q, Liu X, Lin W, Wang T, Wei J, Zhang Z, Lu X, Wu J, Ni D, Song K, Zhang L, Lu X, Lu S, Zhang J. *Nucleic Acids Res*, 2023, 51: W33–W38
- 63 Pillai A, Idris A, Philomin A, Weidle C, Skotheim R, Leung PJY, Broerman A, Demakis C, Borst AJ, Praetorius F, Baker D. *Nature*, 2024, 632: 911–920
- 64 Lu S, Li S, Zhang J. *Med Res Rev*, 2014, 34: 1242–1285
- 65 Lu S, Zhang J. *J Med Chem*, 2019, 62: 24–45

Research progress on design and discovery of allosteric drugs

Rui Li^{1,2,4}, Xinheng He^{1,4}, Lanxin Long¹, Jian Zhang^{1,2,3*}

¹ Department of Pharmaceutical and Artificial-Intelligence Sciences, Shanghai Jiao Tong University School of Medicine, Shanghai 200025, China

² State Key Laboratory of Medical Genomics, National Research Center for Translational Medicine at Shanghai, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai 200025, China

³ Key Laboratory of Protection, Development, and Utilization of Medicinal Resources in Liupanshan Area, Ministry of Education, Peptides & Protein Drug Research Center, School of Pharmacy, Ningxia Medical University, Yinchuan 750004, China

⁴ State Key Laboratory of Drug Research, Shanghai Institute of Materia Medica, Chinese Academy of Sciences, Shanghai 201203, China

*Corresponding author (email: Jian.zhang@sjtu.edu.cn)

Abstract: The identification of allosteric binding sites constitutes a pivotal relationship between structural biology and computational biology, significantly improving the advancement of allosteric drug discovery. Nevertheless, the predominant studies in allosteric drug development heavily depend on high-throughput screening methodologies, which are blocked attributable to the insufficient elucidation of allosteric mechanisms. This review includes insights derived from case studies on allosteric mechanisms, protein structural databases, and the evolution of computational algorithms, with the perspectives of increasing our understanding of allostery and generating more efficacious allosteric drug development. A fundamental component in this domain is the cooperation of structural biology with computational biology, which is indispensable for the translation of three-dimensional structural datasets into actionable drug discovery insights. These datasets, in application with artificial intelligence algorithms, highlight the identification of allosteric binding sites, thereby facilitating the establishment of structure-activity relationships. Furthermore, they are catalyzing the development of computational algorithms specifically designed for allosteric proteins, thus accelerating the progression of the allosteric drug discovery field.

Keywords: allostery effect, artificial intelligence algorithm, drug discovery and development

doi: [10.1360/SSC-2025-0046](https://doi.org/10.1360/SSC-2025-0046)